Review Paper on Object Tracking

Shafali Jaiswal[1] Shailesh Chaurasia[2] Vidushi Chaudhary[3] Deependra Sharma[4]
[4]Assistant Professor
[1,2,3,4]Department of Instrumentation & Control
[1,2,3,4]JSS Academy of Technical Education Noida, India

Abstract— In object recognition the main steps used are the detection of moving object, its tracking and identification. Object detection and tracking are important and challenging tasks in many computer vision applications such as surveillance, vehicle navigation, and autonomous robot navigation. Detection of moving objects in video streams is the first relevant step of information and background subtraction is a very popular approach for foreground segmentation. It is a challenge to understand the most suitable method of tracking and detection algorithm. There is a variety of object tracking, detection and identification algorithms available along with its publications. This paper proposes a systematic review of these algorithms.

Key words: Object recognition, object detection, object identification, surveillance, vehicle navigation, autonomous robot navigation, background subtraction, foreground segmentation

I. INTRODUCTION

This paper proposes research conducted so far for object detection, tracking and identification of objects in video surveillance system. The surveillance system is the process of monitoring the behaviour, activities or the other changing information, for the purpose of influencing, managing, directing, and protecting. This paper outlines the set of challenges discussed in several domains of research work and the majority of relevant work will be reviewed. Tracking is the process of determining the object of interest within a sequence of frames, from its first appearance to its last one.[1] The type of object and its description within the system depends on the application. Object tracking systems are typically geared towards surveillance application where it is desired to monitor people or vehicles moving about an area. The first process in the motion detection is capturing the image information using a video camera.

The motion detection stage includes some image pre-processing step such as; gray-scaling and smoothing, reducing image resolution using low resolution image technique, frame difference, morphological operation and labelling. The pre-processing steps are applied to reduce the image noise in order to achieve a higher accuracy of the tracking. The smoothing technique is performed by using median filter. The low resolution image is performed in three successive frames to remove the small or fake motion in the background. Then frame difference is performed on those frames to detect the moving object emerging in the scene. The next process is applying morphological operation such as dilation and erosion as filtering to reduce the noise that is remained in the moving object. Connected component labelling is then performed to label each moving object in different label. The second stage is tracking the moving object. In this stage, we perform a block matching technique to track only the interest moving object among the moving objects emerging in the background. The blocks are defined by dividing the image frame into non-overlapping square parts. The blocks are made based on PISC image that considers the brightness change in all the pixels of the blocks relative to the considered pixel.

The last stage is object identification. For this purpose we use spatial and color information of the tracked object as the image feature. Then, a feature queue is created to save the features of the moving objects. When the new objects appear on the scene, they will be tracked and labelled, and the features of the object are extracted and recorded into the queue. Once a moving object is detected, the system will extract the features of the object and identify it from the identified objects in the queue.

There are two district approaches to the tracking problem, top-down and another one is bottom-up. Top-down methods are goal oriented and therefore bulk of tracking systems are designed in this manner. It typically involve some sort of segmentation to locate region of interest, from which objects and features can be extracted for the tracking. In this research paper various background subtraction techniques available in the literature word analysed. Background subtraction involves the absolute difference between the current image and the reference updated

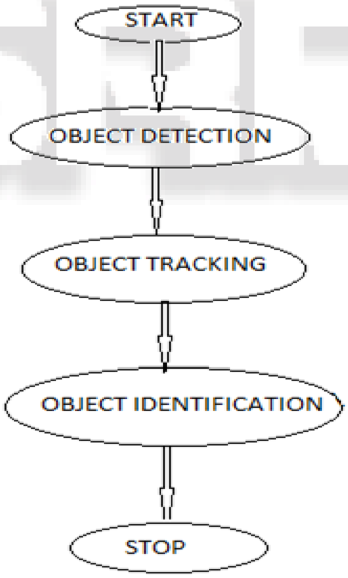

Fig. 1: Entire flow of procedure.

background over a period of time. The problem of varying illumination condition, background clutter, shadows, camouflage, bootstrapping can be overcome with the help of a good background subtraction. In the presence of variability illumination condition, background motion, complex object shape, partial and full object occlusions, object tracking becomes a very challenging task. For it we need to first detect the object and segment it from the video images and then track it from across different frames. The result of detection is used as initialization procedure for tracking[22]. Object tracking is a process that tracks the root of the object over time by locating its location in every frame of the video sequences.

II. LITERATURE SURVEY

An object is an entity of interest. It can be represented by their shapes and appearances. So there are various representations of object shape, which is commonly used for tracking. The joint shape and appearance representations is described as follows[1]:-

A. Points:

The point representation is suitable when it is given more concentration on the object which occupies small regions in an image.

B. Primitive Geometric Shape:

Primitive geometric shapes are more suitable for representing simple rigid objects as well as non-rigid objects.

C. Object Silhouette and Contour:

The region inside the contour is called the silhouette of the object. These representations are suitable for tracking complex non rigid shapes.

D. Articulated Shape Models:

Articulated objects are composed of body parts that are held together by joints.

E. Skeletal Model:

Object skeleton can be extracted by applying medial axis transform to the object silhouette. This model is commonly used as a shape representation for recognizing objects.
Object representations. (a) Centroid, (b) multiple points, (c) rectangular patch.

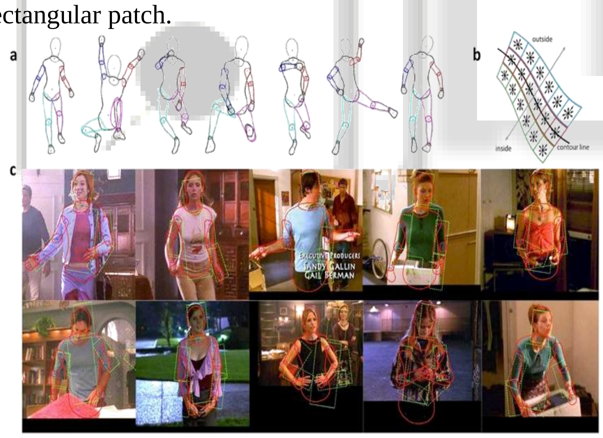

Fig. 2: Object representation.

Similarly the appearance feature of objects can be represented in various other ways. The shape representation can be combined with appearance representation for tracking. Some common appearance representations in the case of object tracking are described as follows[1]:-

F. Probability Densities of Object Appearance:

The probability density estimates the object appearance can either be parameters, such as Gaussian and a mixture of Gaussians, such as Parzen windows and histograms. The probability densities of object appearance features (color, texture) can be computed from the image regions specified by the shape models.

G. Templates:

Templates are formed using simple geometric shapes or silhouettes. It carries both spatial and appearance information. Templates, however, only encode the object appearance generated from a single view. Thus, they are only suitable for tracking objects whose poses do not vary considerably during the course of tracking.

H. Active Appearance Models:

These are generated by simultaneously modelling the object shape and appearance. Object shape is defined by a set of landmarks. Each landmark, an appearance vector is stored in the form of color, texture or gradient magnitude. These models required a training phase where both the shapes & its associated appearance is learned from a set of samples.

I. Multi View Appearance Models:

These models encode different views of an object. One approach to represent the different object views is to generate a subspace from the given view. One limitation of multi-view appearance models is that the appearances in all views have required a lot of time.

In general terms, there is a strong relationship between object representation and tracking algorithms.

III. OBJECT DETECTION AND TRACKING METHODS

Every tracking method requires an object detection mechanism either in every frame or when the object first appears in the video. A common approach for object detection is to use information in a single frame. However, some object detection methods make use of the temporal information computed from a sequence of frames to reduce the number of false detections. The aim of an object tracker is to generate the trajectory of an object over time by locating its position in every frame of the video. But tracking has two definition one is in literally it is locating a moving object or multiple object over a period of time using a camera.

A. Background Subtraction Method:

In video processing applications, variants of the background subtraction method are broadly used for the detection of moving objects in video sequences. The background subtraction [20] is the most popular and common approach for motion detection. Object detection can be achieved by building a representation of the scene called the background model and then finding deviations from the model for each incoming frame. Any significant change in an image region from the background model signifies a moving object. The pixels constituting the regions undergoing change are marked for further processing. This process is referred to as the background subtraction. Unfortunately, a simple inter-frame difference with global threshold reveals itself as being sensitive to basic assumptions of background subtraction. These assumptions are based on a camera which is firmly fixed and has a static noise-free background[1] whereas real-life systems have camera jitters, illumination changes and etc.[4] Absolute difference is taken between every current image It(x; y) and the reference background image B(x; y) to find out the motion detection mask *D*(*x; y*) in simple background subtraction.

$$D(x;y) = \begin{cases} 1; \text{ if } |It(x;y) - B(x;y)| \geq t \\ 0; \text{ otherwise} \end{cases} \quad (1)$$

Where t is a threshold, which decides whether the pixel is foreground or background [2]. If the absolute difference is greater than or equal to t, the pixel is classified as foreground, otherwise the pixel is classified as background[4].

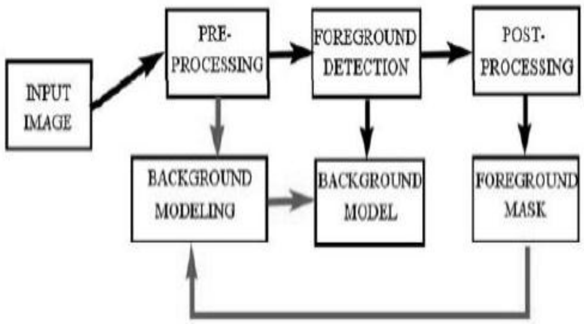

Fig. 3: Background subtraction based on object detection and tracking system architecture.

In literature there are various background subtraction techniques which are gaussian mixture model, kernel density estimation, inter-frame minimum, maximum difference and etc.[2]. The problem with background subtraction [18] is to automatically update the background from the incoming video frame. The four main steps of background subtraction are pre-processing, background modelling, foreground detection, and data validation[3].

a) Gaussian Mixture Model:

The use of Gaussian mixture model is essential for detecting moving objects in video surveillance system. This model has the colour values of particular pixel as a mixture of Gaussians. Foreground is the pixel values that do not fit the background distributions Stauffer and Grimson have proposed a probabilistic approach using a mixture of Gaussian for identifying the background and foreground objects. The probability of observing a given pixel value P_t at time t is given by:-

$$P(p_t) = \sum w_{i,t}\; n(p_t, \mu_{i,t}, \textstyle\sum_{i,t}) \quad (2)$$

The normalized Gaussian n is a function of $w_{i,t}$, $\mu_{i,t}$, $\sum$ i, t which represents weight, mean and co-variance matrix of the ith Gaussian at time respectively. The gaussian mixture can be represented as shown below, depending upon the match.

$$w_{i,t} = (1-\alpha)w_{i,t-1} + \alpha \quad (3)$$

$$\mu_{i,t} = (1-\rho)\mu_{t-1} + p_t\rho \quad (4)$$

$$\sigma^2_{i,t} = (1-\rho)\sigma^2_{i,t-1} + \rho_t - \mu_t)^T (p_t - \mu_t) \quad (5)$$

$$\rho = \alpha\eta\,(\rho_t \,|\mu_{i,t-1}, \sigma_{i,t-1}) \quad (6)$$

In this case the variable defines the speed at which the distribution parameter changes. If the pixel (p_t) matches the i-th Gaussian, then the, matches remaining are updated in the following manner

$$w_{i,t} = (1-\alpha)w_{i,t-1} \quad (7)$$

$$\mu_{i,t} = \mu_{i,t-1} \quad (8)$$

$$\sigma^2_{i,t} = \sigma^2_{i,t-1} \quad (9)$$

Then (b) distribution are modelled to be the background and the remaining (k-b) distributions are modelled as the foreground for the next pixel.

The values for B is determined:-

$$B = argmin_b(\textstyle\sum w_i > T) \quad (10)$$

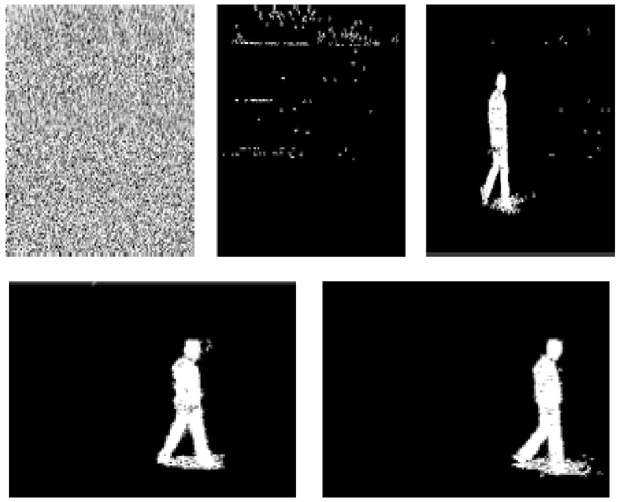

Fig. 4: Gaussian mixture

There are some foreground detection approaches. The most common foreground detection approach is to check whether the input pixel is significantly different from the corresponding background estimate:-

$$|I_t(x,y) - B_t(x,y)| > \tau \quad (11)$$

Where $I_t(x, y)$ and are used to denote the luminance pixel intensity and its background estimate at spatial location (x, y) and time t. equation number 12 shows another popular foreground detection scheme is to apply a threshold based on the normalized statistics[4] :-

$$\frac{|I_t(x,y) - B_t(x,y) - \mu_d|}{\sigma_d} > \tau_s \quad (12)$$

In the above equation μ_d and σ_d are the mean and the standard deviation of $I_t(x, y)-B_t(x,y)$ for all spatial locations (x,y). Fuentes and Velastin proposed one possible modification for threshold determination:-

$$\frac{|I_t(x,y) - B_t(x,y)|}{B_t(x,y)} > \tau_s \quad (13)$$

Where τ_c is used to denote the contrast threshold. There are some other approaches for background subtraction which represents the intensity variations of a pixel in an image sequence. It uses a discrete states corresponding to the events in the environment. Hidden Markov Models[5,6] are used to classify small blocks of an image under background state, foreground state and shadow state. This model successful for certain events, which are hard to model correctly using unsupervised background modelling approaches.

B. Running Average:

Simple background subtraction cannot handle illumination variation and results in noise in the motion detection mask. The problem of noise can be overcome, if the background is made adaptive to temporal changes and updated in every frame.

$$B_t(x; y) = (1 - \alpha)\, B_{t-1}(x; y) + \alpha I_t(x; y) \quad (14)$$

Where B_t is a learning rate. The binary motion detection mask D(x, y) is calculated as follows:-

$$D(x,y) = \begin{cases} 1, \text{ if } |I_t(x,y) - B(x,y)| \geq \tau \\ 0\text{ , otherwise} \end{cases} \quad (15)$$

Fig. 5: Original video frame.

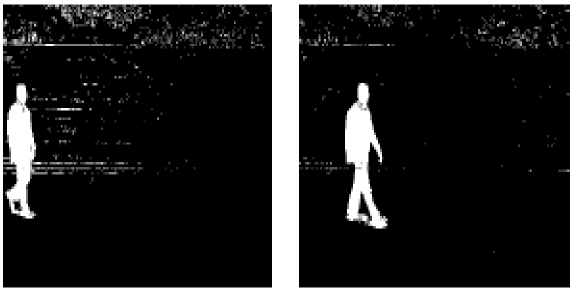

Fig. 6: Output after background subtraction

Fig. 7: Two moving object Fig 8. Single moving Object

IV. COLOR FEATURE EXTRACTION

The color feature extracted from the object is RGB color space as the RGB color information can be obtained from video capture device directly. We extract the information from upper

Fig. 9: Output after background subtraction

and lower part of the object to obtain more colour information for identification. The first color information calculated is mean value of each human body part as calculated for upper part for lower part. The mean value is calculated for each color component of RGB space[22].

$$\mu_{f_k}^{O_U^i} = \frac{\sum_{x=x_{min}^i}^{x_{max}^i} \sum_{y=y_{min}^i}^{y_{max}^i} f_k(x,y)}{\#O_U^i} \quad (16)$$

$$\mu_{f_k}^{O_L^i} = \frac{\sum_{x=x_{min}^i}^{x_{max}^i} \sum_{y=y_{min}^i}^{y_{max}^i} f_k(x,y)}{\#O_L^i} \quad (17)$$

where *i* is number of the moving objects and (x, y) is the coordinate of pixels in moving object. (x^i_{max} ,y^i_{max}) and (x^i_{max} ,y^i_{max}), are the coordinates of the bounding box of moving object *i*, $f_k(x,y)$ denotes pixel value for each color component in RGB space of the current frame, O^i_U and O^i_L denote the set of coordinates of upper and lower part of human body of moving object *i* and *#Oi* is the number of pixels of moving object *i*. The standard deviation has proven to be an extremely useful measure of spread in part because it is mathematically tractable. Standard deviation is a statistical term that provides a good indication of volatility. It measures how widely the values are dispersed from the average. Dispersion is the difference between the actual value and the average value. The larger the difference between the actual color and the average color is, the higher the standard deviation will be, and the higher the volatility.

V. OBJECT TRACKING BASED ON PARTICLE FILTER ALGORITHM

Object tracking can be extremely complex and time consuming especially when it is done in outdoor environments. Some problems is countered while object tracking in outdoor environments are fake-motion background, illumination changes, shadows and presence of clutter. There are variety of tracking algorithms have been proposed and implemented to overcome these difficulties. They can be divided into two categories : deterministic methods and stochastic methods[22].

Deterministic methods perform iterative search for a similarity between the template image and the current one for tracking an object. This algorithm is utilised by the following methods are: Background subtraction, in terframed difference, optical flow, skin color extraction and so on. Whereas the stochastic methods use the state space to model the underlying dynamics of the tracking system such as Kalman filter and particle filter. Kalman filter is a common approach for dealing with target tracking in the probabilistic framework. It can be used to propagate and update the covariance and mean of the distribution of this model. But it cannot resolve the tracking problem when the model is nonlinear and non-Gaussian. To overcome those problems, particle filter has been introduced by many researchers and become popular algorithm to estimate the problem of nonlinear and non-Gaussian estimation framework. The particle filter, also known as sequential Monte Carlo is the most popular approach which recursively constructs the posterior pdf of the state space using Monte Carlo integration. It approximates a posterior probability density of the state such as the object position by using samples or particles. Although particle filters have been widely used in recent years, they have important drawback such as samples are spread around several modes pointing out the different hypotheses in the state space. Many improvements have been introduced, but there is still much ground to cover. Different approaches have been taken in order to overcome these problems used a particle filter based on color histograms features. Histograms are robust to partial occlusions and rotations but no shape analysis is taken into account. However, the problem for object tracking with color occurs when the region around the object is cluttered and illumination is change. In this way, a color feature based tracking does not provide reliable performance because it fails to fully model the target especially when occlusion occurs. In another work, approach relied on gradient-based optimization and color-based histograms. In this case, no dynamic model is used therefore no occlusion can be predicted presented an interesting approach called annealing particle filter which aims to reduce the required number of samples. However, it could be inappropriate in a cluttered environment. They combine edge and intensity measures but they focused on motion analysis, and thus, no occlusion handling is explored. To overcome the above problems, in this article, we made some improvements on color based object tracking and employed it to track a moving object We proposed an object state not only object position but also speed, size, object size scale and appearance condition of the object.

VI. MORPHOLOGICAL OPERATIONS

Morphological operation is performed to fill small gaps inside the moving object and to reduce the noise remained in the moving objects. The morphological operators implemented are dilation followed by erosion. The process of dilation adds pixels to the boundary of the object and closes isolated background pixel. Dilation of set A by structuring element B [7] is defined as:-

A ® B= $U(A)_b$ (18)

The number of pixels added or the number of pixels removed from the objects in an image depends on the size and shape of the structuring element used to process the image. Morphological operation eliminates background noise and fills small gaps inside an object. There is no fixed limit on the number of times dilation and erosion is performed. In the given algorithm dilation and erosion is used iteratively till the foreground object is completely segmented from the background. After morphological operation now the results of following frames, remove noise from frame difference and background subtraction frame result.

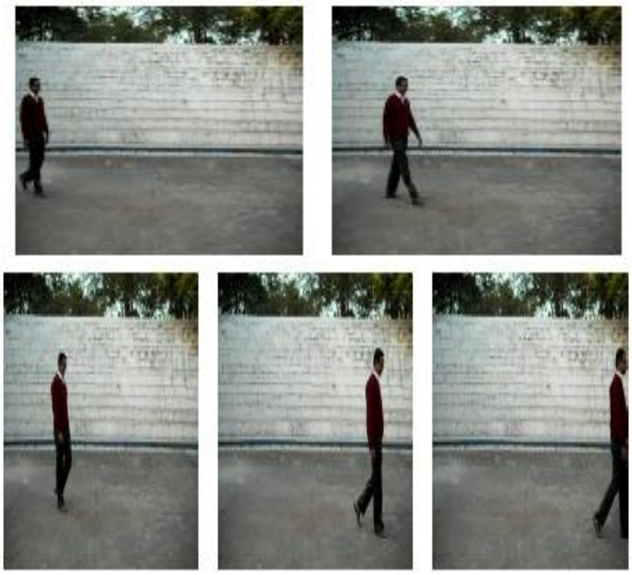

Fig. 10: Original video frames

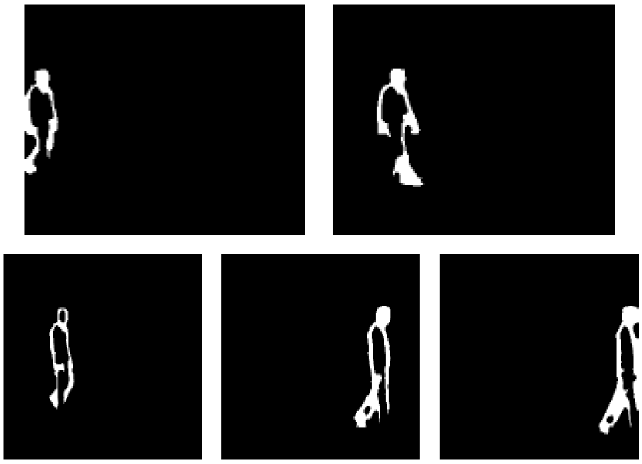

Fig. 11: Output of frame difference result after noise removal

VII. BACKGROUND MODELLING

The basic principle of background subtraction is to compare a static background frame with the current frame of the video scene pixel by pixel. This technique builds a model of the background any frame can be compared with the model to detect zones where a significant difference occurs. Some steps for the background subtraction, i.e. first developement of a background model of the scene is done, then background subtraction to detect foreground object.

Sugandi et al proposed tracking technique of moving persons using camera peripheral increment sign correlation image [10].Julio Cezar[19] has proposed a background model, and incorporates a novel technique for shadow detection in gray scale video sequences . In the first stage, a pixel wise median filter over time is applied to several seconds of videos to distinguish moving pixels from stationary pixels.

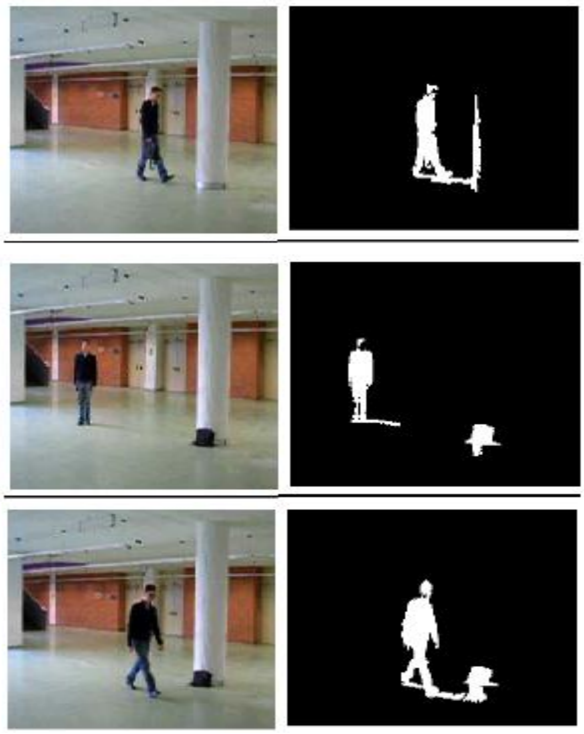

Fig. 12: Output after background subtraction using Julio Cezar method

VIII. ACTIVE CONTOUR MODEL

Active contour model also known as ACMs, uses an object tracking process it is in scope of edge-based segmentation. A snake which is an energy minimizing spline, is a type of active contour model. Its energy is based on its shape and location within the image. A connected to each other and can easily be deformed under applied force. Total energy of a snake is defined by the equation given below[21], where γ is the external energy weight. A snake can be assumed as a group of control points connected to each other and can easily be deformed under applied force. The situation, in which a snake works the most abundant is the situation where the points are at the adequate distance and the situation[25], in which the initial position's coordinates are controlled. Total energy of a snake is defined as [25] in equation 19:-

$$E_{snake}=\frac{1}{2}\int_s[\alpha(s)|\upsilon_s|^2+\beta(s)|\upsilon_{ss}|^2]+\gamma E_{external}(\upsilon(s))ds \qquad (19)$$

where,

$$\upsilon_s=\frac{d\upsilon(s)}{ds} \qquad (20)$$

and

$$\upsilon_{ss}=\frac{d^2\upsilon(s)}{ds^2} \qquad (21)$$

IX. REGION-BASED METHOD

When a moving object is segmented, a region of pixels assigned to the object is available which can be tracked using approaches like cross-correlation. A moving object consists of one or more tracked regions. Combination of several regions into one object is then performed at a higher level of abstraction.

There are several techniques available for modelling and tracking image regions. The regions are often modelled using a probability density distribution of their colour. This distribution can be described using a colour histogram, or a mixture of Gaussian kernels. Instead of using one 3D probability density distribution, separate distributions for each of the colours can be used.

Probability density distributions is not a dependent variable of changes in object orientation, scale, partial occlusion, viewing position and object deformation. This makes this method interesting for tracking non-rigid objects such as human being. However, the distributions capture Region-Based Moving Object Detection and Tracking only, the colours in an image and do not include any spatial correlation information. Therefore, they have limited discriminative power. A color correlogram, on the other hand, is a co-occurrence matrix that gives the probability that a pixel at a distance d from a given pixel of colour is of colour [12]. This way spatial information in the form of distance to pixels of a certain colour is introduced.

Other approach is by taking spatial information into account by using many small regions and using the time average per-pixel colour . Instead of choosing one colour space, automatic selection of most discriminative features can be used . This adapts the colour space used, which is done by comparing the specific tracked region with the local background, leading to more precise object segmentation. However, when pixels are misclassified and consequently used for updating the wrong model, this solution becomes unstable.

Moving objects can also be modelled using a fixed or parameterized shape, like in the mean-shift approach [13], and the particle filter [14]. However such techniques have many disadvantages which is that they are unable to describe an arbitrary shape, changing between subsequent frames.

X. THE FEATURE-BASED APPROACH

This method finds features, for example, image edges, corners, and other structures well localized in two dimensions and tracks these as they move from one frame to another. This involves two steps:

Firstly, the features are found in two or more consecutive images. The act of feature extraction will both reduce the amount of information to be processed, and also go some way towards obtaining a higher level of understanding of the case, by eliminating the unimportant parts. Secondly, these features are matched between the frames. In the simplest and commonest case, two frames are used and two sets of features are matched to give a single set of motion vectors. Alternatively, the features in one frame can be used as seed points at which to use other methods, for example, gradient-based methods to find the flow. The two stages of feature-based flow estimation each have their own problems. The feature detection stage requires features to be located accurately and reliably. This has proved to be a non-trivial task, and much work has been carried out on feature detectors[15]. The feature matching stage has the problem of occurrence of ambiguous potential matches; unless image displacement is smaller than the distance between the features.

XI. OBJECT IDENTIFICATION

This is the last step of anomalous detection where we identify the object which is exceeding the threshold value defined by the user. Feature extraction is the main unit of the object identification which is further classified into two types:- color and spatial information of the moving object.

A. Spatial Feature Extraction:

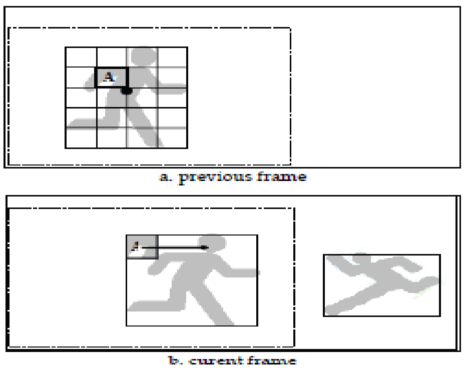

Fig. 13: Matching process

The feature of objects extracted in the spatial domain is the position of the tracked object. The trajectory defines the spatial information combined with the features of the object in the time domain. Frome this the movement and speed of the moving object can be estimated.

After we obtain the interesting object, it is extracted by using the boundary box. The bounding box can be determined by computing the maximum and minimum value of x and y coordinates of interesting moving object according to the following equation:-

$$B^{i}_{min} = \{(x^{i}_{min}, y^{i}_{min})|x,y\epsilon O^{i}\} \quad (22)$$

$$B^{i}_{max}=\{(x^{i}_{max}, y^{i}_{max})|x,y\epsilon O^{i}\} \quad (23)$$

where O^i denotes the set of the coordinate of points in the interest moving object i, B^i_{min} is the left-top corner coordinates of the interest moving object *i* and B^i_{max} is the right-bottom corner coordinates of the interest moving object *i*, respectively.

B. Color Feature Extraction:

The video captured provides the RGB color information directly, and the color feature extracted from this is the RGB color space.

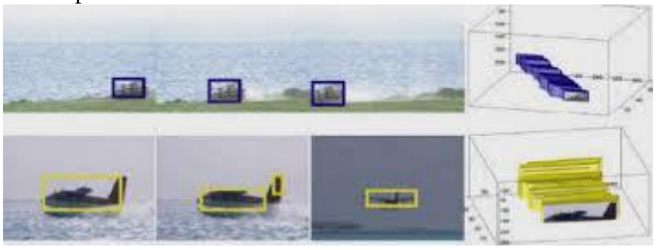

Fig. 14: Example of bounding box of moving object

Good feature should have following properties:-

1) Distinctiveness:

The features should be such that it should be highly distinguishable and matched.

2) Repeatability:

If we have two images taken under different viewing conditions of the same objector view, then a high sense of the observable part should be observed in both the images.

3) Locality:

To decrease the probability of occlusion and to permit simple model estimates of the geometric and photometric deformations between two images taken under different viewing circumstances the features should be limited.

4) Quantity:

To detect a sensible number of features on small objects, the number of detected features should be satisfactorily large.

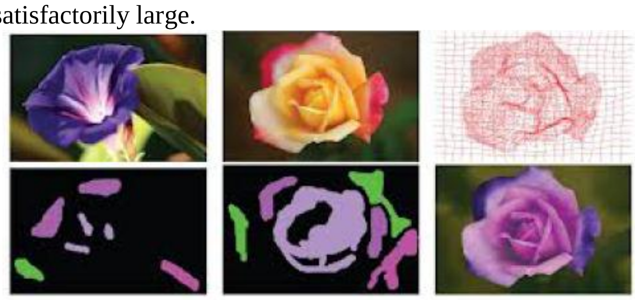

Fig. 15: Image features

XII. CONCLUSION

In this paper we presented the commonly implemented object tracking, detection and identification techniques. After studying all of it, a comparative study was conducted to observe that no method outperforms the other one on each video category. More research, however, is needed to improve the robustness against the effects of the environment such as noise, illumination changes, occlusions and etc. We require to question about the weaknesses of particular algorithm against specific conditions. The aim of this paper is to provide a better understanding performances of video surveillance systems in the literature via published measures, computational and environmental details.

REFERENCES

[1] Hanna Goszczynska, “On object tracking", published by InTech, 2011.

[2] Benezeth, Y., Jodoin, P.M., Emile, B., Laurent, H., and Rosenberger, C, "Review and evaluation of commonly-implemented background subtraction algorithms", International Conference on Pattern Recognition, 2008.

[3] Cheung, S.-C., Kamath, "On robust techniques for background subtraction in urban traffic video", Video Communications and Image Processing, SPIE Electronic Imaging, January 2004.

[4] Yilmaz, A., Javed, O., Shah, M., “Object tracking: A survey”, ACM Computing Surveys, Vol. 38, Dec. 2006.

[5] Rittscher, J., Kato, J., Joga, S., and Blake, A, "A probabilistic background model for tracking", European Conference on Computer Vision (ECCV), Vol. 2, 2002.

[6] Stenger, B., Ramesh, V., Paragios, N., Coetzee, F., and Buhmann, "On topology free hidden Markov models: Application to background modelling', IEEE International Conference on Computer Vision (ICCV), 2002.

[7] Muhammad Shoaib, Ralf Dragon, and Jorn Ostermann,, "" On shadow detection for moving humans using gradient-based background subtraction", IEEE International Conference on, pages 773–776, 2009.

[8] Intan Kartika and Shahrizat Shaik Mohamed, " On Frame differencing with post-processing techniques for moving object detection in outdoor environment", In Signal Processing and its Applications (CSPA), 2011 IEEE 7th International Colloquium on, pages 172–176. IEEE, 2011.

[9] Gouet-Brunet, V., Lameyre, "On object recognition and segmentation in videos by connecting heterogeneous visual features", Computer Vision and Image Understanding, Vol. 111, Special Issue on Intelligent Visual Surveillance (IEEE), 2008.

[10] Mahbub Murshed1/2, Md Hasanul Kabir1/2, and Oksam Chae1/2,"'Moving object tracking-an edge segment based approach", 2011.

[11] Jianhua Ye, Tao Gao, and Jun Zhang., "Moving object detection with background subtraction and shadow removal", Fuzzy Systems and Knowledge Discovery (FSKD), 2012 9th International Conference on, pages 1859–1863. IEEE, 2012.

[12] M.B. Capellades, David S. Doermann, Daniel DeMenthon, and RamaChellappa, "An appearance based approach for human and object tracking", IEEE International Conference on Image Processing (ICIP), 2008.

[13] Dorin Comaniciu, Visvanathan Ramesh, and Peter Meer, "Real–time tracking of non– rigid objects using mean shift", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2000.

[14] Patrick P´erez, Carine Hue, Jaco Vermaak, and Michel Gangnet, "On Color–based probabilistic tracking", the

European Conference on Computer Vision (ECCV), 2002.
[15] H. Asada and M. Brady", The curvature primal sketch", IEEE Trans. on Pattern Analysis and Machine Intelligence,1986.
[16] Zhiming Qian, Hongxing Shi , Jiakuan Yang and Lianxin Duan, "Video-based multiclass vehicle detection and tracking", IJCSI International Journal of Computer Science Issues, January 2013.
[17] Zhan Chaohui, Duan Xiaohui, Xu Shuoyu, Song Zheng, and Luo Min, "On an improved moving object detection algorithm based on frame difference and edge detection", Fourth International Conference on, pages 519–523. IEEE, 2007.
[18] Cezar Silveira Jacques, Claudio Rosito Jung, and Soraia Raupp Musse , " On Background subtraction and shadow detection in gray scale video sequences", Computer Graphics and Image Processing, 2005. SIBGRAPI 2005. 18th Brazilian Symposium on, pages 189–196. IEEE, 2005.
[19] Budi Sugandi, Hyoungseop Kim, Joo Kooi Tan, and Seiji Ishikawa, 2008, "A block matching technique for object tracking employing peripheral increment sign correlation image", Computer and Communication Engineering ICCCE 2008. International Conference on, pages 113–117.IEEE, 2008.
[20] Lloyd L Coulter, Douglas A Stow, Yu Hsin Tsai, Christopher M Chavis, Richard W Mc Creight, Christopher D Lippitt, and Grant W Fraley, 2, "On A new paradigm for persistent wide area surveillance", Homeland Security (HST), IEEE Conference on Technologies for, pages 51–60. IEEE, 2012.
[21] Alper Yilmaz, Omar Javed, Mubarak Shah, " On object tracking: A survey", University of Central Florida, 2006.
[22] Dagher, I., Tom, K. E., "WaterBalloons, A hybrid watershed Balloon Snake segmentation", Image and Vision Computing, 2008.